AI glossary for conference attendees

AIPH 2026

A list of terms to help attendees follow the presentations and workshops at AIPH 2026

Term	Definition
AI agent	An AI system that uses a model to decide and carry out steps towards a goal, using tools where needed. For example, it might search outbreak reports, extract details and prepare a briefing, within defined permissions. It has an LLM at it's heart and a degree of autonomy to determine next steps and schedule tasks and respond to inputs.
AI governance	The responsibilities, rules and processes used to decide whether and how AI should be developed, purchased, used and monitored, including accountability for errors and harms.
Algorithm	A set of computational rules or steps for completing a task. In machine learning, an algorithm is used to learn patterns from data.
Algorithmic bias	Systematic differences in an AI system's outputs or performance that can unfairly disadvantage particular groups. Bias can arise from data, design choices or how the system is used.
Application programming interface (API)	A defined way for software systems to communicate automatically. For example, a surveillance program can send text to an AI model through an API and receive its analysis.
Artificial intelligence (AI)	Computer systems that perform tasks such as recognising patterns, interpreting language, generating content or making predictions and recommendations.
Cross-validation	Evaluating a modelling approach by repeatedly training and checking it on different portions of a dataset. This helps assess consistency of performance; a separate test dataset may still be needed.
Data sovereignty / Indigenous Data Sovereignty	Data sovereignty concerns authority over data and their use. Indigenous Data Sovereignty specifically recognises Indigenous peoples' rights to govern data about their peoples, communities, lands and resources.
Deep learning / neural network	A neural network is a machine-learning model made of connected processing units. Deep learning uses networks with many layers to learn complex patterns, such as those in medical images.
Distillation / knowledge distillation	Training a "student" model to reproduce aspects of a "teacher" model's behaviour, using its outputs or other training signals. Often used to create smaller, faster or cheaper models, although performance and errors can also be transferred.

Term	Definition
Embedding / semantic search	An embedding represents information as numbers that capture aspects of its meaning. Semantic search uses these representations to find related content, even when different words are used.
Ensemble model	A method that combines predictions from several models to improve performance. Random forests and gradient boosting are examples; an ensemble can also combine different model types.
Explainable AI / explainability	Methods that help people understand which information influenced an AI output. An explanation can support scrutiny, but does not itself establish that the output is correct.
F1 score	A measure combining precision and recall using their harmonic mean. It ranges from 0 to 1; higher values indicate a better balance between avoiding false alerts and missing genuine events.
Federated learning	Training a shared model across organisations while keeping the underlying records locally. Participants share model updates rather than pooling patient records; additional privacy protections may still be needed.
Fine-tuning	Further training an existing model on selected examples to adapt it to a particular task, such as extracting cancer diagnoses from pathology reports. This changes the model itself.
Generative AI (GenAI)	AI that creates text, images, audio or other content using patterns learned from training data. Its outputs can be useful without being factually correct.
Gradient boosting	A machine-learning method that builds models sequentially, with each helping correct earlier errors. XGBoost, LightGBM and CatBoost are widely used implementations based on decision trees.
Ground truth / reference standard	The answers used as the benchmark when training or evaluating a model, such as expert-reviewed diagnoses. This benchmark can itself contain errors or disagreements.
Hallucination	AI-generated information that is false, fabricated or unsupported by the source material, often presented convincingly. Examples include invented references, incorrect case counts or nonexistent guidelines.
Harness (AI or agent harness)	The software surrounding an AI model that manages its instructions, information, tools and workflow. A harness can enable the model to search the web, work with files and complete tasks in several steps.
Human-in-the-loop	A workflow in which people review, correct or approve AI outputs at defined stages—for example, checking automatically assigned injury codes before accepting them into a surveillance dataset.
Information extraction / entity extraction	Automatically identifying specific details within text, such as diseases, locations, dates and case counts, and organising them into usable fields.

Term	Definition
Large language model (LLM)	An AI model trained on large amounts of text to process and generate language. LLMs power many conversational tools and can summarise, translate, extract information and draft text. Many current models also process images or audio.
Local model	Most interaction with large language models occur on the internet via powerful computers in data centres. It is also possible to install smaller models on local computers with powerful GPUs. They usually have limited intelligence and hallucinate more than frontier cloud models. (Small models require computers that cost $5-12k, for near frontier performance requires >$200k).
Machine learning (ML)	A type of AI that learns patterns from data to make predictions or classifications, rather than requiring a person to specify every rule.
Model	A trained computational system that learns patterns from data. Claude, Gemini and Grok name both model families and the assistants built around them. ChatGPT is an application using OpenAI models, including the GPT family. These are examples of LLMs or large language models within a proprietary harness.
Model training / training data	Training is the process through which a model learns patterns from examples. Training data are those examples; their quality and representativeness influence how the model performs.
Model validation / testing	Checking how well a model performs, especially on data it did not learn from. External validation checks performance in another population, organisation or setting.
Multi-agent system	A system in which several AI agents share work or perform different roles—for example, searching for evidence, analysing findings and checking a report. They may use the same underlying model or different models.
Multimodal AI	AI that works with more than one type of information, such as text, images, audio or numerical data. For example, it may combine medical images with patient records.
Natural language processing (NLP)	Computer methods for working with human language, such as identifying injury details in clinical notes or classifying news reports about outbreaks.
Open-source intelligence (OSINT)	Information gathered and analysed from publicly accessible sources, such as news, websites and social media. In public health, it can help identify emerging threats. Here, “open-source” refers to information sources, rather than software licensing.
Optical character recognition (OCR)	Technology that converts text in images or scanned documents into machine-readable text—for example, extracting words from photographs of advertisements.
Overfitting	When a model learns the peculiarities of its training data so closely that it performs poorly on new data.

Term	Definition
Precision	Of everything the model identifies as positive, the proportion that is actually positive. For outbreak surveillance: “Of the alerts raised, how many were genuine events?” Also called positive predictive value.
Prompt / prompt engineering	A prompt is the instruction, question and supporting information given to an AI tool. Prompt engineering means designing and refining these inputs to improve its responses.
Random forest	A machine-learning method that combines many decision trees—each using branching questions—to produce a prediction or classification.
Recall (sensitivity)	Of all the actual positives, the proportion the model successfully identifies. For outbreak surveillance: “Of the genuine events, how many did the system detect?”
Retrieval-augmented generation (RAG)	A method that retrieves relevant information from documents or databases and supplies it to an AI model when it answers. It can improve relevance but does not guarantee accuracy or safety.
Skill / SKILL.md file	A reusable set of instructions telling an AI agent how to perform a particular task. SKILL.md is the main instruction file, written in Markdown, a simple text format. Skills may include references, templates and scripts; they guide the agent without retraining its model.
Structured / unstructured data	Structured data have predefined fields, such as age and diagnosis in a spreadsheet. Unstructured data include free-text clinical notes, images and recordings. AI can help extract structured information from them.
Supervised learning	Training a model using examples with known answers—for example, pathology reports already labelled by experts as showing cancer or no cancer.
Synthetic data	Artificially generated data designed to resemble real data, such as simulated patient records. They can support development and testing, but require checks for realism, bias and privacy.
Token	A unit of text processed by a language model, such as a word, part of a word or punctuation. Input limits and usage costs are often measured in tokens.
Zero-shot prompting	Asking an AI model to perform a task without giving it worked examples in the prompt—for example, asking it to classify a news article using written instructions alone.

Prepared by craig.dalton@newcastle.edu.au – please provide feedback or suggest revisions.

Definition references

NSW Health Agency for Clinical Innovation — https://aci.health.nsw.gov.au/ai-framework/glossary

Google Machine Learning Glossary — https://developers.google.com/machine-learning/glossary/

NIST retrieval-augmented generation — https://csrc.nist.gov/glossary/term/retrieval_augmented_generation

Anthropic agent harnesses — https://claude.com/resources/articles/harnessing-claudes-intelligence

Agent Skills specification — https://agentskills.io/specification

IBM knowledge distillation — https://www.ibm.com/think/topics/knowledge-distillation